

RECEIVED

NOV 20 2002

TECH CENTER 1600/2900



1600

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/724,676

DATE: 11/13/2002

TIME: 16:33:51

#10

Input Set : D:\pat.ST25.txt

Output Set: N:\CRF4\11132002\I724676.raw

ENTERED

ENTERED

```

3 <110> APPLICANT: Compugen LTD
5 <120> TITLE OF INVENTION: Variants of alternative splicing
7 <130> FILE REFERENCE: 129181.4 Compugen
9 <140> CURRENT APPLICATION NUMBER: 09/724,676
10 <141> CURRENT FILING DATE: 2000-11-28
12 <160> NUMBER OF SEQ ID NOS: 97222
14 <170> SOFTWARE: PatentIn version 3.2
16 <210> SEQ ID NO: 1
17 <211> LENGTH: 745
18 <212> TYPE: DNA
19 <213> ORGANISM: Homo sapiens
21 <400> SEQUENCE: 1
22 cggggctgcc cttcttccct gcggaggagg ggcctgggag gtcgcgttgg cgggagggag      60
24 gttacctttc ccagtctcgc tctggccgcc tgagccagga ggaagcagcg gcgaggtctg      120
26 cgggagggcat ggcgggagct ccggacgagc gccggcgggg ccccgcgga ggggagcagc      180
28 tgcagcagca acacgtctct tgccaggtct tccccgagcg tctggcccag gggaatcccc      240
30 agcaagggtt cttctccagc ttcttcacca gcaaccagaa gtgccagctt aggtcctga      300
32 agacgctgga gacaagtagg agccatgacc tggaggctgt ggttccacag aatgggtctg      360
34 agaccgggtg ggcgaggaag ggcctgggca acacctggcc aggggcttca ggttcagcac      420
36 agagtctaga cagactgggc atcatggggg ctgggttggg agcctgaggc tcagtttccc      480
38 cactagaggt ggcgctgaga cccttttctg cgtgtttttt ggcttgtgga atgaatttgt      540
40 ctgccctttt ttgtcactca ctgttgtgac agatcctccg atttcttaca taagtattgc      600
42 agccacactg tgacttcagg gtaagggtga ggccaagttt taaagtgcc gccagggttaa      660
44 ttaaggcaca caaaggctgt aaaaatgtaa aaatgtaatt aaggcatacc aaggctgtaa      720
46 aaatataaaa gcagttttgc aaacc
49 <210> SEQ ID NO: 2
50 <211> LENGTH: 754
51 <212> TYPE: DNA
52 <213> ORGANISM: Homo sapiens
54 <220> FEATURE:
55 <221> NAME/KEY: misc_feature
56 <222> LOCATION: (14)..(14)
57 <223> OTHER INFORMATION: n is a,c,g, or t
59 <220> FEATURE:
60 <221> NAME/KEY: misc_feature
61 <222> LOCATION: (23)..(23)
62 <223> OTHER INFORMATION: n is a,c,g, or t
64 <220> FEATURE:
65 <221> NAME/KEY: misc_feature
66 <222> LOCATION: (25)..(26)
67 <223> OTHER INFORMATION: n is a,c,g, or t
69 <400> SEQUENCE: 2
70 tgcattggaat ttgnacgaga ccntnntgaa ccaccggcct gtgactgctg tggagatcca      60

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/724,676

DATE: 11/13/2002

TIME: 16:33:51

Input Set : D:\pat.ST25.txt

Output Set: N:\CRF4\11132002\I724676.raw

```

72 gctgatggtg gaagagagtg aagagcggct cacggaggag cagattgaag ctcttctcca 120
74 caccgtcacc agcattctgc ctgcagagcc agaggctgag cagaagaaga atacaaacag 180
76 caatgtggca atggacgaag aggacccagc atagaagagc acagctggcc cggcggtttc 240
78 atgaagtcag aaggcctggc agccatttcc tggacgttga gaggattgtt tatttgattt 300
80 ttatcctcat cccagcaggc ctggctttgt ggtagtttg gtacatcaca aaaataagtt 360
82 aaaaagaaat atttgtgcct tggggagaag aaacatggtg aaaacaggct gaggttgta 420
84 gggcagagag ctgaagggtg ggacagtgc cgcggacccc tctgtgcttg aaagatttcc 480
86 tccacggcct ttgccccagt tgtggggagg tctctgtgca cagcggggaa aatgcttggt 540
88 tcgccttttg tgggccatgt cctaattagt ttcattctgt tccctgggaa cttactaagg 600
90 ggcccagagc actgttgga gtctggttag agtccccaga gatttactct aagttaaaat 660
92 gagccactga ccttggtcct ccttagagga atttcctcga gaacaacaga gataagaaaa 720
94 gaaccggcct ggccaatcct tcaacagctc taga 754
97 <210> SEQ ID NO: 3
98 <211> LENGTH: 1292
99 <212> TYPE: DNA
100 <213> ORGANISM: Homo sapiens
102 <220> FEATURE:
103 <221> NAME/KEY: misc_feature
104 <222> LOCATION: (1244)..(1244)
105 <223> OTHER INFORMATION: n is a,c,g, or t
107 <220> FEATURE:
108 <221> NAME/KEY: misc_feature
109 <222> LOCATION: (1251)..(1251)
110 <223> OTHER INFORMATION: n is a,c,g, or t
112 <400> SEQUENCE: 3
113 cggcacgagg gagatgatga gttgctgagt gtgcacacct ttccggaaca catacacaca 60
115 cctgctctg ggatcccttg tgaggctgcc ctcatggagt tccccctggc ccagatatgt 120
117 ccccaagcac ctccagatc acagacatga cccgcaggag ctgccagaac ctgggctaca 180
119 ctgcggcatc tccccaggcc ccggaggctg cctccagcac aggggaatgct gagagggcag 240
121 aggaggtgcc tggagaagga agcctgttcc tgcaggccga gaccggggt tggttccaga 300
123 agaccaggc cactggctc ctgcagcac gggcagcccc tgctgtgttc catggcttca 360
125 tcaccgggag ggaggcagag aggtgctgg agcccaagcc tcagggggtgc tacttggtgc 420
127 ggttcagcga gagcgcggtg accttcgtgc tgacttacag gagccggact tgctgccgcc 480
129 acttcctgct ggcccagctc agggacgggc gccacgtggt gctgggcgag gacagcgccc 540
131 acgcgcggct gcaggacctg ctgctgcaat acaccgcgca cccgctcagc ccctacgggg 600
133 agacgctcac cgagccctc gcccgacaga ctctgagcc tgcaggactt tccctgagga 660
135 ccgaagaatc aaactttgga agcaaaagcc aggacccaaa cccccagta agcccaatca 720
137 tcaaacaggg gcaagcccca gtcccgatgc agaaagagg ggccggggag aaggagccct 780
139 cccagctgct caggcccaag cctcccatcc ccgcaaaacc tcagctgccc ccagaagtct 840
141 acacaatccc tgttccacga caccgcccgg ccccaagccc caagccctcc aatcctatct 900
143 acaatgagcc tgatgaaccc atagctttct atgccatggg ccggggcagc cctggggaag 960
145 cccccagcaa catctatgtg gaagtggaag atgaggcct acccgccacc cttgggcacc 1020
147 ctgtcctacg gaagagctgg tccaggcctg tcccaggagg ccagaatata ggtggctccc 1080
149 agctgcattc tgagaactct gtgattgggc aaggccctcc cctgccccac cagccccac 1140
151 ccgctggag acacaccctc cccacaatc tttctagaca ggtgcttcag gacagaggac 1200
153 aggcattggt tccccttggg cctcctcagt aggcggtaag tctnagcttt ngtgaggata 1260
155 actttggagc atcaggttcc agagacctt ca 1292
158 <210> SEQ ID NO: 4
159 <211> LENGTH: 1323

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/724,676

DATE: 11/13/2002

TIME: 16:33:51

Input Set : D:\pat.ST25.txt

Output Set: N:\CRF4\11132002\I724676.raw

160 <212> TYPE: DNA  
 161 <213> ORGANISM: Homo sapiens  
 163 <220> FEATURE:  
 164 <221> NAME/KEY: misc\_feature  
 165 <222> LOCATION: (1275)..(1275)  
 166 <223> OTHER INFORMATION: n is a,c,g, or t  
 168 <220> FEATURE:  
 169 <221> NAME/KEY: misc\_feature  
 170 <222> LOCATION: (1282)..(1282)  
 171 <223> OTHER INFORMATION: n is a,c,g, or t  
 173 <400> SEQUENCE: 4

174	cggcacgagg	gagatgatga	gttgctgagt	gtgcacacct	ttccggaaca	catacacaca	60
176	ccctgctctg	ggatcccttg	tgaggctgcc	ctcatggagt	tccccctggc	ccagatatgt	120
178	ccccaaagga	gtcacgaagc	ccccatccca	accttcagca	ccttccagat	cacagacatg	180
180	accgcgagga	gctgccagaa	cctgggctac	actgcggcat	ctccccaggc	cccggaggct	240
182	gcctccagca	cagggaatgc	tgagagggca	gaggagggtg	ctggagaagg	aagcctgttc	300
184	ctgcaggccg	agaccggggc	ttggttccag	aagaccaggg	cccactggct	cctgcagcac	360
186	ggggcagccc	ctgcctgggt	ccatggcttc	atcaccggga	gggaggcaga	gaggctgctg	420
188	gagcccaagc	ctcaggggtg	ctacttggtg	cggttcagcg	agagcgcggt	gaccttcgtg	480
190	ctgacttaca	ggagccggac	ttgtgcccgc	cacttcctgc	tggcccagct	cagggacggg	540
192	cgccacgtgg	tgctggggca	ggacagcgcc	caagcgcggc	tgcaggacct	gctgctgcac	600
194	tacaccgcgc	accgcctcag	cccctacggg	gagacgctca	ccgagcccct	cgcccgacag	660
196	actcctgagc	ctgcaggact	ttccctgagg	accgaagaat	caaactttgg	aagcaaaagc	720
198	caggacccaa	acccccagta	cagcccaatc	atcaaacagg	ggcaagcccc	agtcccgatg	780
200	cagaaagagg	gggcccggga	gaaggagccc	tcccagctgc	tcaggcccaa	gcctcccatc	840
202	cccgcctaac	ctcagctgcc	cccagaagtc	tacacaatcc	ctgttccacg	acaccgcccg	900
204	gccccacgcc	caaagccctc	caatcctatc	tacaatgagc	ctgatgaacc	catagctttc	960
206	tatgccatgg	gccggggcag	ccctggggaa	gccccagca	acatctatgt	ggaagtggaa	1020
208	gatgagggcc	taccgcgccac	ccttggggcac	cctgtcctac	ggaagagctg	gtccaggcct	1080
210	gtcccaggag	gccagaatac	aggtggctcc	cagctgcatt	ctgagaactc	tgtgattggg	1140
212	caaggccctc	ccctgccccca	ccagccccca	cccgcctgga	gacacaccct	cccccaaat	1200
214	ctttctagac	aggtgcttca	ggacagagga	caggcatggc	ttccccttgg	gcctcctcag	1260
216	taggcggtaa	gtctnagctt	tngtgaggat	aactttggag	catcaggttc	cagagacctt	1320
218	tca						1323

221 <210> SEQ ID NO: 5  
 222 <211> LENGTH: 2119  
 223 <212> TYPE: DNA  
 224 <213> ORGANISM: Homo sapiens  
 226 <400> SEQUENCE: 5

227	accgccacca	tgaacaagct	ttacatcggc	aacctcaacg	agagcgtgac	ccccgcggac	60
229	ttggagaaa	tgtttgcgga	gcacaagatc	tcctacagcg	gccagttctt	ggtcaaatcc	120
231	ggctacgcct	tcgtggactg	cccggacgag	cactgggcga	tgaaggccat	cgaaactttc	180
233	tccgggaaa	tagaattaca	aggaaaacgc	ttagagattg	aacattcggt	gccccaaaaa	240
235	caaaggagcc	ggaaaattca	aatccgaaat	attccacccc	agctccgatg	ggaagtactg	300
237	gacagcctgc	tggtcagta	tggtacagta	gagaactgtg	agcaagtga	caccgagagt	360
239	gagacggcag	tggtgaatgt	cacctattcc	aaccgggagc	agaccaggca	agccatcatg	420
241	aagctgaatg	gccaccagtt	ggagaacccat	gccctgaagg	tctcctacat	ccccgatgag	480
243	cagatagcac	agggacctga	gaatggggcg	cgagggggct	ttggctctcg	gggtcagccc	540
245	cgccaggctc	accttgtggc	agcggggggc	cagccaagca	gcagcaagtg	gacatcccc	600

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/724,676

DATE: 11/13/2002

TIME: 16:33:51

Input Set : D:\pat.ST25.txt

Output Set: N:\CRF4\11132002\I724676.raw

247	ttcggctcct	ggtgcccacc	cagtatgtgg	gtgccattat	tggcaaggag	ggggccacca	660
249	tccgcaacat	cacaaaacag	accaggtcca	agatagacgt	gcataggaag	gagaacgcag	720
251	gtgcagctga	aaaagccatc	agtgtgcact	ccaccctga	gggctgctcc	tccgcttgta	780
253	agatgatctt	ggagattatg	cataaagagg	ctaaggacac	caaaacggct	gacgagggttc	840
255	ccctgaagat	cctggcccat	aataactttg	tagggcgtct	cattggcaag	gaaggacgga	900
257	acctgaagaa	ggtagagcaa	gataccgaga	caaaaatcac	catctcctcg	ttgcaagacc	960
259	ttacccttta	caaccctgag	aggaccatca	ctgtgaagg	ggccatcgag	aattgttgca	1020
261	gggccgagca	ggaaataatg	aagaaagtcc	gggaggccta	tgagaatgat	gtggctgcc	1080
263	tgagcctgca	gtctcacctg	atccctggcc	tgaacctggc	tgctgtaggt	cttttccag	1140
265	cttcatccag	cgcagtcocg	ccgcctccca	gcagcgttac	tggggctgct	ccctatagct	1200
267	cctttatgca	ggctcccgag	caggagatgg	tgcaggtgtt	tatccccgcc	caggcagtgg	1260
269	gcgccatcat	cggcaagaag	gggcagcaca	tcaaacagct	ctcccggttt	gccagcgct	1320
271	ccatcaagat	tgcaccaccc	gaaacacctg	actccaaagt	tcgtatgggt	atcatcactg	1380
273	gaccgccaga	ggcccaattc	aagggtctcc	agtatgatag	ctaagcagaa	ggaaagctga	1440
275	caagctgccc	acttgagcct	tgcccttcac	ctgccttaga	tggctcaggg	aagaatctat	1500
277	ggcaaactca	aggaggagaa	cttctttgtg	cccaaggagg	aagtgaagct	ggagaccac	1560
279	atacgtgtgc	cagcatcagc	agctggccgg	gtcattggca	aagggtgaaa	aacggtgaac	1620
281	gagttgcaga	atttgacggc	agctgaggtg	gtagtaccaa	gagaccagac	ccctgatgag	1680
283	aacgaccagg	tcacgtgaa	aatcatcgga	catttctatg	ccagtcagat	ggctcaacgg	1740
285	aagatccgag	acatccctggc	ccagggttaag	cagcagaaat	cgagagtgtg	ctctccccgg	1800
287	caggcctgag	aatgagtggg	aatccgggac	acctggggccg	ggctgtagat	caggtttgcc	1860
289	cacttgattg	agaaagatgt	tccagtgagg	aaccctgac	tctcagcccc	aaacacccac	1920
291	ccaattggcc	caacactgtc	tgccctcgg	ggtgtcagaa	attctagcgc	aaggcacttt	1980
293	taaactgtga	ttgtttaaag	aagctctcca	ggccccacca	agagggtgga	tcacacctca	2040
295	gtgggaagaa	aaataaaatt	tccttcaggt	tttaaaaaca	tgcagagagg	tgttttaatc	2100
297	agccttaag	gatggttca					2119
300	<210>	SEQ ID NO: 6					
301	<211>	LENGTH: 2040					
302	<212>	TYPE: DNA					
303	<213>	ORGANISM: Homo sapiens					
305	<400>	SEQUENCE: 6					
306	accgccacca	tgaacaagct	ttacatcggc	aacctcaacg	agagcgtgac	ccccgcggac	60
308	ttggagaaag	tgtttgcgga	gcacaagatc	tctacagcg	gccagttctt	ggtcaaatcc	120
310	ggctacgcct	tcgtggaactg	cccggacgag	cactgggcga	tgaaggccat	cgaaactttc	180
312	tccgggaaag	tagaattaca	aggaaaacgc	ttagagattg	aacattcggt	gccccaaaaa	240
314	caaaggagcc	ggaaaattca	aatccgaaat	attccacccc	agctccgatg	ggaagtactg	300
316	gacagcctgc	tggtcagta	tggtacagta	gagaactgtg	agcaagtga	caccgagagt	360
318	gagacggcag	tggtgaatgt	cacctattcc	aaccgggagc	agaccaggca	agccatcatg	420
320	aagctgaatg	gccaccagtt	ggagaacat	gacctgaagg	tctctacat	ccccgatgag	480
322	cagatagcac	agggacctga	gaatggggcg	cgagggggct	ttggctctcg	gggtcagccc	540
324	cgccaggctc	accttgtggc	agcggggggc	cagccaagca	gcagcaagtg	gacatcccc	600
326	ttcggctcct	ggtgcccacc	cagtatgtgg	gtgccattat	tggcaaggag	ggggccacca	660
328	tccgcaacat	cacaaaacag	accaggtcca	agatagacgt	gcataggaag	gagaacgcag	720
330	gtgcagctga	aaaagccatc	agtgtgcact	ccaccctga	gggctgctcc	tccgcttgta	780
332	agatgatctt	ggagattatg	cataaagagg	ctaaggacac	caaaacggct	gacgagggttc	840
334	ccctgaagat	cctggcccat	aataactttg	tagggcgtct	cattggcaag	gaaggacgga	900
336	acctgaagaa	ggtagagcaa	gataccgaga	caaaaatcac	catctcctcg	ttgcaagacc	960
338	ttacccttta	caaccctgag	aggaccatca	ctgtgaagg	ggccatcgag	aattgttgca	1020
340	gggccgagca	ggaaataatg	aagaaagtcc	gggaggccta	tgagaatgat	gtggctgcc	1080

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/724,676

DATE: 11/13/2002

TIME: 16:33:51

Input Set : D:\pat.ST25.txt

Output Set: N:\CRF4\11132002\I724676.raw

```

342 tgagcctgca gtctcacctg atccctggcc tgaacctggc tgctgtaggt cttttcccag 1140
344 cttcatccag cgcagtcccg ccgcctccca gcagcggttac tggggctgct ccctatagct 1200
346 cctttatgca ggctcccag caggagatgg tgcaggtggt tatccccgcc caggcagtg 1260
348 gcgccatcat cggcaagaag gggcagcaca tcaaacagct ctcccggttt gccagcgct 1320
350 ccatcaagat tgcaccacc gaaacacctg actccaaagt tcgtatggtt atcatcactg 1380
352 gaccgccaga ggcccaattc aaggctcagg gaagaatcta tggcaaaactc aaggaggaga 1440
354 acttcttttg tcccaaggag gaagtgaagc tggagacca catacgtgtg ccagcatcag 1500
356 cagctggccg ggtcattggc aaaggtggaa aaacggtgaa cgagttgcag aatttgacgg 1560
358 cagctgaggt ggtagtacca agagaccaga cccctgatga gaacgaccag gtcacgtga 1620
360 aaatcatcgg acatttctat gccagtcaga tggctcaacg gaagatccga gacatcctgg 1680
362 cccaggttaa gcagcagaaa tcgagagtgt gctctccccg gcaggcctga gaatgagtgg 1740
364 gaatccggga cacctgggccc gggctgtaga tcaggtttgc ccacttgatt gagaaagatg 1800
366 ttccagtggg gaacctgat ctctcagccc caaacacca cccaattggc ccaacactgt 1860
368 ctgcccctcg ggggtgtcaga aattctagcg caaggcactt ttaaactggg attgtttaaa 1920
370 gaagctctcc agggccccc accaggggtgg atcacacctc agtgggaaga aaaataaaat 1980
372 ttccctcagg ttttaaaaac atgcagagag gtgttttaaat cagccttaaa ggatggttca 2040
375 <210> SEQ ID NO: 7
376 <211> LENGTH: 1990
377 <212> TYPE: DNA
378 <213> ORGANISM: Homo sapiens
380 <400> SEQUENCE: 7
381 accgccacca tgaacaagct ttacatcggc aacctcaacg agagcgtgac cccgcggac 60
383 ttggagaaag tgtttgcgga gcacaagatc tctacacgag gccagttctt ggtcaaatcc 120
385 ggctacgcct tcgtggactg cccggacgag cactgggcga tgaaggccat cgaaacttct 180
387 tccgggaaag tagaattaca aggaaaacgc ttagagattg aacattcggg gccccaaaaa 240
389 caaaggagcc ggaaaattca aatccgaaat attccacccc agctccgatg ggaagtactg 300
391 gacagcctgc tggtcagta tggtacagta gagaactgtg agcaagtga caccgagagt 360
393 gagacggcag tggtgaatgt cacctattcc aaccgggagc agaccaggca agccatcatg 420
395 aagctgaatg gccaccagtt ggagaacccat gccctgaagg tctcctacat ccccgatgag 480
397 cagatagcac agggacctga gaatgggcgc cgagggggct ttggtctctg ggtcagccc 540
399 cgcaggctc accctgtggc agcggggggc cagccaagca gcagcaagtg gacatcccc 600
401 ttcggtcctt ggtgcccacc cagtatgtgt gtgccattat tggcaaggag ggggccacca 660
403 tccgcaacat cacaaaacag acccagtcga agatagacgt gcataggaag gagaacgcag 720
405 gtgcagctga aaaagccatc agtgtgcact ccaccctga gggctgtctc tccgcttgta 780
407 agatgatctt ggagattatg cataaagagg ctaaggacac caaaacggct gacgaggttc 840
409 ccctgaagat cctggcccat aataactttg tagggcgtct cattggcaag gaaggacgga 900
411 acctgaagaa ggtagagcaa gataccgaga caaaaatcac catctcctcg ttgcaagacc 960
413 ttacccttta caaccctgag aggaccatca ctgtgaaggg ggccatcgag aattgttgca 1020
415 gggccgagca ggaaataatg aagaaagtcc gggaggccta tgagaatgat gtggtgcca 1080
417 tgagcctgca gtctcacctg atccctggcc tgaacctggc tgctgtaggt cttttcccag 1140
419 cttcatccag cgcagtcccg ccgcctccca gcagcggttac tggggctgct ccctatagct 1200
421 cctttatgca ggctcccag caggagatgg tgcaggtggt tatccccgcc caggcagtg 1260
423 gcgccatcat cggcaagaag gggcagcaca tcaaacagct ctcccggttt gccagcgct 1320
425 ccatcaagat tgcaccacc gaaacacctg actccaaagt tcgtatggtt atcatcactg 1380
427 gaccgccaga ggcccaattc aaggctcagg gaagaatcta tggcaaaactc aaggaggaga 1440
429 acttcttttg tcccaaggag gaagtgaagc tggagacca catacgtgtg ccagcatcag 1500
431 cagctggccg ggtcattggc aaaggtggaa aaacggtgaa cgagttgcag aatttgacgg 1560
433 cagctgaggt ggtagtacca agagaccaga cccctgatga gaacgaccag gtcacgtga 1620
435 aaatcatcgg acatttctat gccagtcaga tggctcaacg gaagatccga gacatcctgg 1680

```